About Segmentation Step in Content-based Image Retrieval Systems

Jérôme da Rugna, Chareyron Gael, Konik Hubert

About Segmentation Step in Content-based Image Retrieval Systems

Jérôme Da Rugna, Gael Chareyron and Hubert Konik

***Abstract*—Despite of the hope arised a few years ago, Content Based Image Retrieval - CBIR - systems has not reached the initial goal, ie to manage and search images in database: we are unable to link the semantic sens of an image to numerical values. However, some members of the community have begun the necessary introspection. The analyze of each step of the feature extraction will allow us to overcome actual problematics and to take the right path in the future. In this context, we propose in this paper to discuss about a low-level tool frequently used: the segmentation step. In the general context of scene images, we evaluate the stability of some classical algorithms using a basic protocol. The quite inefficiency of all approaches let us conclude to the necessity to use meta-data and any other collected informations during this first segmentation step.**

***Index Terms*—Content-based image retrieval, segmentation, evaluation.**

I. Introduction

TO browse and query multimedia object, and specially images, in a database is one of the highest goal of the last decennary[1], [2], [3], [4], [5], [6], [7], [8], [9]. However, if some years ago the publication amount was very important, since the beginning of this millenary, the community seems globally less interested by this topic, even if, at the opposite, the industry still asks for systems, even more with the growth of private numerical images databases. Of course, the poor results obtained by Content Based Image Retrieval systems were confronted to the user expectancy and, right now, the main majority still use the old archiving method based on keywords.

The community of image retrieval by content using image processing algorithm is then actually separated in three parts. First, the negative mind : it does not work and we will never reach the generic image search by content. Secondly, the optimist mind, it is not so bad and we just have to continue in the same direction. In this case, many optimization or variant of existing algorithms are proposed but without erasing the problem, that is to say the final user's dissatisfaction. The third way, the pragmatic way, is to consider that time is now to answer major questions as

- "Why it is not working as required ?";
- "How improving the features extraction ?";
- "What can we expect from the future approaches ?".

To answer these questions in a terms of image processing, it needs foremost to make a concrete evaluation on all contributor tools in the image search path. In this paper, we will then discuss the segmentation step in the feature extraction chain, step frequently used in CBIR systems. For example, if this step is insufficient enough, the process depends on fragile foundations...

Jérôme Da Rugna and Gael Chareyron are with Laboratory GRID-SILV, ESILV, Paris La Defense, Francee-mail:gael.chareyron@devinci.fr, jerome.da_rugna@devinci.fr

Hubert Konik is with Laboratory Hubert Curien, University of Saint-Etienne, France

II. The segmentation debate

Many systems use segmentation to calculate image features[10], [11], [12]. In fact, the segmentation step is necessary to extract information from images. Indeed, how is it possible to find an object in an image without first extracting its own regions ? Several kinds of segmented region may be extracted according to the homogeneity predicates used: color, texture or else semantic criteria. Many approaches exist and many of them propose an adapted and optimized version of some well-known classical algorithms or other.

Even if a segmentation step is required in a feature extraction process, the expected results from this low level process still have to be discussed. Of course, if the segmentation was able to extract semantic regions, the discussion would be close. However, this is not the case and we can not expect from the future such an algorithm because of the underlying diversity. If we consider the semantic extraction impossible, we have to first estimate the robustness of this segmentation step in order to really select the suitable parameters during the feature extraction.

Nevertheless, at this point of research in image retrieval by content, the question of using or not using segmentation may be asked. And, in this case, which method or approach may be used. Whatever these interrogations are often inquired, few objective studies permit to answer it. Our goal in this paper is then to propose an objective evaluation of the stability of classical image segmentation process. Our goal is not to set up a new evaluation protocol but this paper must be considered as a ponctual step forward to overcome the use of segmentation algorithms in image retrieval by content. In the context of natural scene images, we will first merely measure the stability of some algorithms made for feature extraction, without judge the quality of segmentation itself.

III. Evaluate segmentation stability

A. Evaluation: an image retrieval point of view

Most of evaluation techniques of segmentation results measure one of the next two properties: ([13], [14], [15],[16], [17], [18]).

- The difference from a reference segmentation, obtained generally manually;
- The segmentation quality, evaluated by intrinsic relation between segmented regions.

Nevertheless, these approaches are not adapted to our goal of segmentation evaluation in an image retrieval by content oriented context. The first way supposes the existence of a perfect segmentation of an image, which is really difficult in an exhaustive point of view. To catch the segmentation

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

quality by evaluating intra and inter entropy is interesting in many contexts. But, criteria like the Borsotti[14] measure, do not yield enough information to the image retrieval segmentation problem. It does not need to measure the quality of region extraction but to measure how the segmentation step reach the expected results from the feature extraction chain. Then, we purpose to evaluate segmentation by two measures.

- Mix coefficient
  Measure the capacity of the segmentation process to extract an object from a background, ie if the segmentation process mixes the object with the background in a same region, it would be even more difficult to retrieve a posteriori this object.
- Superposition rate
  Measure the region extraction spatial stability of the segmentation process, ie, instead of evaluating the segmentation result, first already judge the capacity of the method to always extract the same spatial regions, independently of a background.

B. Evaluation protocol

1) Overall schema: Let first precise the protocol used for the evaluation. The figure 1 shows the overall schema.

Let explain the different components.

- The object
  An object is chosen for the entire evaluation. The results is then object oriented and will, a priori, be dependent on the object complexity.

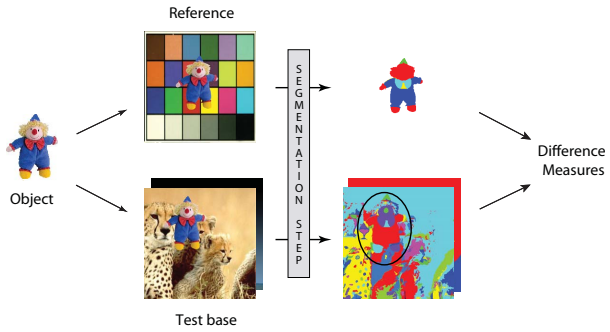

Fig. 1. Overall schema of the evaluation process

- The reference
  In order to extract a reference segmentation, we insert an object in a reference image and segment this image, as shown in figure 2. Finally we extract obtained regions. The choice of the reference image, ie a color chart, will of course impact the results. Anyway, our experiences with many different backgrounds have shown that the few modification obtained permit, in this discussion, to retain only one background.

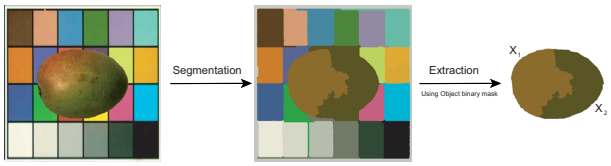

Fig. 2. Reference segmentation creation

- The test base
  A set of 3000 images of natural scenes of indoor or outdoor pictures is used. This set is enough numerous to obtain measure stability. Then the object is inserted in each image at a random position, as illustrated in the figure 3. The position is stored, not to be used during the segmentation step, but to be able to compute the two considered following measures.

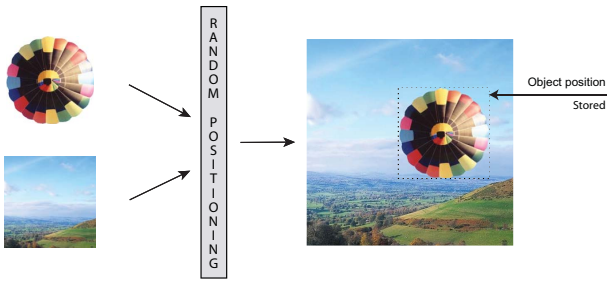

Fig. 3. Object insertion in scene images

2) MC: Mix Coefficient: Let $Y = \bigcup_{j=1}^{n} Y_j$ a segmentation of the object X. $\bar{X}$ the complementary part, ie the part of the image not recovered by the object X. Then the Mix Coefficient is defined by :

$$MC = \frac{\sum_{j=1}^{n} (Card(Y_j \cap X) \times \delta_j)}{Card(X)} \quad (1)$$

with

$$\delta_j = \begin{cases} 1 \text{ If } \dfrac{Card(Y_j \cap \bar{X})}{Card(Y_j)} \geq t \\ 0 \text{ Else} \end{cases} \quad (2)$$

t is a threshold, set to 5%, that enable a region to have a part of pixels mixed in the background without be considered as mixed.

3) RnS: Rate of no Superposition.: Let $Y = \bigcup_{j=1}^{n} Y_j$ a segmentation of the object X, $X = \bigcup_{j=1}^{m} X_j$ the reference segmentation obtained as shown in the figure 2 and N the size of X. Let G the following bipartite graph :

- $(Y_j)_{1 \leq j \leq n}$ and $(X_j)_{1 \leq j \leq m}$ as vertices
- Edges are set to the surface (in pixels) superposition between regions of Y and regions of X.

Applying a bipartite maximum weighted matching we build then a set C of k couples. C_i is then defined as the number of pixels of the matched regions in X and Y.

Then the Rate of no Superposition is defined by:

$$RnS = \frac{N - \sum_{i=1}^{k} C_i}{N} \quad (3)$$

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

4) Tested segmentation algorithms: We have selected 5 segmentation algorithms, two spatial and three color-based ones. They represent the diversity of all segmentation approaches in CBIR systems. For some kind of approches, we have implemented several specific versions, designed for natural scene images. However, it is noticeable that the results are usually equivalent between image retrieval designed versions and the classical one. Let first describe briefly these methods:

- WaterShed
  Results are shown using the De Andrade and Al version[19].
- Pyramid
  This kind of methods [20], [21] are used to be fast and efficient. We have implemented a linked pyramid, designed for natural images.
- Clustering
  K-means clustering[22] is often considered as a reference. We have implemented for this study the Liew and Al's version [23].
- MeanShift
  We use the wellknown Meanshift algorithm application to image segmentation by Comaniciu and Meer[24].
- Histogram
  To stand for these classical techniques and the many variant[25], [26] we have applied here the hierarchical Cheng's approach[27].

C. First results

First of all, we need to measure the stability of each method according to the context an object based on several colors and textures, like the the parrot picture shown in figure 4. This "object" is also complex considering a shape point of view. We then applied the protocol as previously described and the results are given in the table I.

Fig. 4. The parrot object

The first interpretation of these data is the unexpected poor quality of the results. Let examine separately the two measures:

- The Mix Coefficient is a lot varying from 8%, for Meanshift to 34%, for pyramid. Less than 10% seems a small value, but it means than quite 10% of the object is lost. The figure 5 illustrates what may be 15% and 20% of pixels mixed with the background. Shape retrieval may be very difficult or impossible, even if only 15% of object surface is lost! Moreover, the way the object is inserted in the images, without post-processing for example, must normally make easier the detection as strong contours between object and background are artificially created. Visually at least, it was easy to separate the object from the background.
- The Rate of No Superposition is varying from 32% to 49%. It means that, in average, more than a third of the parrot is not covered by a match between the reference segmentation and the evaluated segmentation. We can note here that we obtained equivalent measures changing the color chart background reference with another one. The high values of this measure show than stability of segmentation algorithms is not sufficient. The partition of the parrot is too strongly influenced by the context in order to be positive regarding the post recognition step.

Globally, these results are always obtained using the same object, without altering it via compression, scale or other perturbation processes. This is yet the best suitable case in an image retrieval process. And even in this simply case, no segmentation approaches would be sufficient enough to recognize the object in each image. It means that, to find a parrot, actual segmentation techniques are unadapted.

TABLE I
RESULTS ON PARROT OBJECT

Method	MC	RnS
WaterShed	0.12	0.34
Pyramid	0.34	0.38
Clustering	0.21	0.49
MeanShift	0.08	0.33
Histogram	0.16	0.44

Fig. 5. Examples where 15% and 20% of pixels are lost in the parrot image

D. Simpler objects

Of course, only the results are presented in this study but we applied our protocol to several objects, chosen according to their different complexity in color, shape or texture. Then, the results are globally the same as in the parrot case. Finally, we propose two new objects, less complex, ie a mango and a lemon. These two objets are shown in figure 6.

Fig. 6. The lemon and the mango examples

The tables II and III show the obtained results. We can conclude that, in a completely surprising way, the results are not improved as regards the parrot object. Nonetheless, results are worst for the Rate of no Superposition. It would say that the complexity of object does not have linear relation with the stability of the segmentation. Some other results, not given here, with other objects confirm this hypothesis.

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE II
RESULTS ON LEMON OBJECT

Method	MC	RnS
WaterShed	0.09	0.39
Pyramid	0.25	0.47
Clustering	0.1	0.15
MeanShift	0.08	0.31
Histogram	0.08	0.28

TABLE III
RESULTS ON MANGO OBJECT

Method	MC	RnS
WaterShed	0.08	0.22
Pyramid	0.45	0.37
Clustering	0.32	0.26
MeanShift	0.02	0.44
Histogram	0.05	0.31

IV. CONCLUSIONS AND FUTURE WORKS

To conclude, lets indicate that the results presented in this study are extracted from a more complex protocol in order to achieve an objective evaluation of the segmentation process in a content-based retrieval approach. Nevertheless, all the data are difficult to be analyzed and we have then chosen to be focused on more significant ones. The implemented segmentation methods give very poor results on an even so easy task, that is to say to extract an object artificially inserted in scene images. The obtained values are so far from the minimal expected results, from what we can built on a post recognition step. Then, how believe into achieving a query image retrieval, which is much more difficult, using segmentation techniques ? An algorithm can not be efficient to segment an image if it is no adapted to this other one. To only know the information (X,Y, Colour) is not enough to interpret the information contained in the image. The retained results given in this study seem to be sufficient to show that it will be not sufficient to simply continue with only proposing new general algorithms in order to improve content-based retrieval. To afford the segmentation, other informations than the image itself have to be linked to the process.

To segment an image, the algorithm must be adapted to the concerned image. Moreover, as each method seems to be more adapted according to the context, why not imagine to create different results with different approaches for each image ? More precisely, our future works will be to construct new oriented-segmentation algorithms which will be able to consider properties like:

- Meta Data
  Many meta information are stored in the image archiving format, like EXIF in JPEG images. A link has to be made between these informations, like the focus or the flash flag, and the intrinsic parameters of the segmentation algorithm.
- Keywords
  If an image is already described by some keywords, these one may be relatively informative to configure parameters of the segmentation.
- Estimable properties
  We are able to estimate directly informations from an image, like “indoor/outdoor” or “blur/No blur”. These computed properties, even if not absolute, may guide the segmentation process.
- What we search
  To know what we search have to be taking account in the region extraction. To search an object or a global impression can not concern the same segmentation algorithm.

We have then listed some future works in order to improve the segmentation process itself but these points suppose the capability to manage global information systems in order to answer favorably to the growing demand.

REFERENCES

[1] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Image databases at the end of the early years,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 1, January 2001.

[2] R. Li, Y. Zhang, Z. Lu, J. Lu, and Y. Tian, “Technique of image retrieval based on multi-label image annotation,” in *Multimedia and Information Technology (MMIT), 2010 Second International Conference on*, vol. 2, april 2010, pp. 10 –13.

[3] C.-H. Chuang, C.-C. Chang, and S.-C. Cheng, “Content aware image segmentation for region-based object retrieval,” in *Proc. International Symposium on Intelligent Signal Processing and Communications Systems ISPACS 2008*, Feb. 8–11, 2009, pp. 1–4.

[4] L. Yang, R. Jin, L. Mummert, R. Sukthankar, A. Goode, B. Zheng, S. Hoi, and M. Satyanarayanan, “A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 1, pp. 30 –44, jan. 2010.

[5] Z. Liang, H. Fu, Z. Chi, and D. Feng, “Image pre-classification based on saliency map for image retrieval,” in *Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on*, dec. 2009, pp. 1 –5.

[6] B. Wang, X. Zhang, M. Wang, and P. Zhao, “Saliency distinguishing and applications to semantics extraction and retrieval of natural image,” in *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on*, vol. 2, july 2010, pp. 802 –807.

[7] S. Zakariya, R. Ali, and N. Ahmad, “Combining visual features of an image at different precision value of unsupervised content based image retrieval,” in *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on*, dec. 2010, pp. 1 –4.

[8] Y. Liu, D. Zhang, G. Lu, and W. Ma, “A survey of content-based image retrieval with high-level semantics,” *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, 2007.

[9] H. Veeraraghavan, P. Schrater, and N. Papanikolopoulos, “Robust target detection and tracking through integration of motion, color, and geometry,” *Computer Vision and Image Understanding*, vol. 103, no. 2, pp. 121–138, 2006.

[10] S. K. Shandilya and N. Singhai, “Article: A survey on: Content based image retrieval systems,” *International Journal of Computer Applications*, vol. 4, no. 2, pp. 22–26, July 2010, published By Foundation of Computer Science.

[11] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, “Content-based multimedia information retrieval,” *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 2, no. 1, pp. 1–19, Feb. 2006. [Online]. Available: http://dx.doi.org/10.1145/1126004.1126005

[12] R. Datta, D. Joshi, J. Li, and J. Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” *ACM Comput. Surv.*, vol. 40, no. 2, pp. 1–60, May 2008. [Online]. Available: http://dx.doi.org/10.1145/1348246.1348248

[13] Y. Zhang, “Evaluation and comparison of different segmentation algorithms,” *Pattern Recognition Letters*, vol. 18, pp. 963–974, 1997.

[14] M. Borsotti, P. Campadelli, and R. Schettini, “Quantitative evaluation of color image segmentation results,” *Pattern Recognition Letters*, vol. 19, pp. 741–747, 1998.

[15] P. Correia and F. Pereira, “Standalone objective segmentation quality evaluation,” *Journal on Applied Signal Processing*, vol. 2002, no. 4, pp. 389–400, 2002.

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

[16] M. Polak, H. Zhang, and M. Pi, "An evaluation metric for image segmentation of multiple objects?" *Image and Vision Computing*, vol. 27, no. 8, pp. 1223–1227, Jul. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.imavis.2008.09.008

[17] J. K. Udupa, V. R. Leblanc, Y. Zhuge, C. Imielinska, H. Schmidt, L. M. Currie, B. E. Hirsch, and J. Woodburn, "A framework for evaluating image segmentation algorithms." *Computerized medical imaging and graphics the official journal of the Computerized Medical Imaging Society*, vol. 30, no. 2, pp. 75–87, 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16584976

[18] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 260 – 280, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1077314207001294

[19] M. de Andrade, G. Bertrand, and A. Araújo, "An attribute-based image segmentation method," *Materials Research*, vol. 2, no. 3, pp. 145–151, 1999.

[20] Y. Deng and B. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 800–810, 2001.

[21] A. D. Bimbo, M. Mugnaini, P. Pala, and F. Turco, "Visual querying by colour perceptive regions," *Pattern Recognition*, vol. 31, pp. 1241–1253, 1998.

[22] R. Duda, P. Hart, and D. Stork, *Pattern Classification (Second Edition)*. Wiley-Interscience, 2001.

[23] A. Liew, S. Leung, and W. Lau, "Fuzzy image clustering incorporating spatial continuity," *Vision, Image and Signal Processing*, vol. 22, pp. 593–601, 2001.

[24] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 24, pp. 438–445, 2001.

[25] P. Lambert and L. Macaire, "Filtering and segmentation: the specifity of colour images," in *CGIP'2000 : 1st International Conference on Color in Graphics and Image Processing*, Saint-Etienne, France, october 2000, pp. 57–71.

[26] Y. I. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," *Computer Graphics and Image Processing*, vol. 13, pp. 222–241, 1980.

[27] H. Cheng, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Trans. on Image Processing*, vol. 9, no. 12, pp. 2071–2082, december 2000.

ISBN: 978-988-18210-9-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)